Development and validation of a radiogenomics model to predict axillary lymph node metastasis in breast cancer integrating MRI with transcriptome data: A multicohort study

Huifang Chen†, Xiaosong Lan†, Tao Yu, Lan Li, Sun Tang, Shuling Liu, Fujie Jiang, Lu Wang, Yao Huang, Ying Cao, Wei Wang, Xiaoxia Wang* and Jiuquan Zhang*

Department of Radiology, Chongqing University Cancer Hospital, Chongqing, China

Introduction: To develop and validate a radiogenomics model for predicting axillary lymph node metastasis (ALNM) in breast cancer compared to a genomics and radiomics model.

Methods: This retrospective study integrated transcriptomic data from The Cancer Genome Atlas with matched MRI data from The Cancer Imaging Archive for the same set of 111 patients with breast cancer, which were used as the training and testing groups. Fifteen patients from one hospital were enrolled as the external validation group. Radiomics features were extracted from dynamic contrast-enhanced (DCE)-MRI of breast cancer, and genomics features were derived from differentially expressed gene analysis of transcriptome data. Boruta was used for genomics and radiomics data dimension reduction and feature selection. Logistic regression was applied to develop genomics, radiomics, and radiogenomics models to predict ALNM. The performance of the three models was assessed by receiver operating characteristic curves and compared by the *Delong* test.

Results: The genomics model was established by nine genomics features, and the radiomics model was established by three radiomics features. The two models showed good discrimination performance in predicting ALNM in breast cancer, with areas under the curves (AUCs) of 0.80, 0.67, and 0.52 for the genomics model and 0.72, 0.68, and 0.71 for the radiomics model in the training, testing and external validation groups, respectively. The radiogenomics model integrated with five genomics features and three

radiomics features had a better performance, with AUCs of 0.84, 0.75, and 0.82 in the three groups, respectively, which was higher than the AUC of the radiomics model in the training group and the genomics model in the external validation group (both $P < 0.05$).

Conclusion: The radiogenomics model combining radiomics features and genomics features improved the performance to predict ALNM in breast cancer.

KEYWORDS

radiogenomics, lymph node metastasis, breast cancer, genomics, radiomics

Introduction

Breast cancer is the most commonly diagnosed cancer among women worldwide and is the second leading cause of cancer-related death (1). Axillary lymph nodes (ALNs) are an important path for lymph node metastasis (LNM) in breast cancer. Axillary lymph node metastasis (ALNM) is an important factor affecting the treatment and prognosis of breast cancer patients. Thus, accurate identification of ALN involvement in patients with breast cancer is essential for prognosis and therapeutic decision-making (2). Sentinel lymph node (SLN) biopsy is now considered as the reference standard for ALN status staging in patients with clinically negative lymph nodes (3). Compared with ALN dissection, it significantly reduces complications such as arm numbness, upper extremity edema, nerve damage, etc. (4, 5). However, the procedure is still invasive. Therefore, a noninvasive and reliable assessment of ALN status pretreatment is critical for clinical decision-making.

Magnetic resonance imaging (MRI) is one of the important auxiliary examination methods widely used for breast cancer patients before surgery due to its high sensitivity for evaluating tumor extension, intraductal spread, and the presence of multicentric or multifocal lesions (6). Radiomics based on MRI as a noninvasive technology has emerged as a potential method in precision medicine for breast cancer (7, 8). Recently, radiomics has been widely used to predict LNM in breast cancer and has demonstrated excellent predictive performance (9–11). Radiomics focuses on the systematic characterization of the aggressiveness of breast cancer by effectively extracting and analyzing massive image data. However, radiomics may lack the ability of tumor characterization in microstructural features.

Previous efforts have been made to explore the identified genomics features, which provide a powerful tool for identifying breast cancer patients with distant recurrence and might provide a better method for individual risk assessment in patients with lymph node-negative breast cancer (12). Other studies revealed

that a miRNA-dependent model could predict LNM in cervical cancer patients (13), and an epigenetic model could predict axillary staging with ER-positive early-stage breast cancer patients (14). The greatest limitation to deploying genome sequencing for clinical application is that tumor spatial heterogeneity limits genomics tissues, and genomics only reflect a microcosm of the genetic code.

Radiogenomics is a rapidly developing method to integrate genomics data with radiomics data (15). Radiogenomics can provide voxel-by-voxel information from genomics to tumor imaging and thereby guide tailored therapy and help to improve our understanding of tumor biology (15). In the earliest radiogenomics study of breast cancer, Yamamoto et al. explored the relationship between MRI features and gene expression (16). A follow-up study that included preoperative DCE-MRI explored the multiscale relationships between DCE-MRI phenotypes, metastasis, and long noncoding RNA expression (17). These results mainly explored the relationships between imaging and genomics. Furthermore, studies have reported that radiogenomics models have a higher prediction performance than genomics-only models in predicting LNM or radiomics-only models in predicting pathologic complete response in triple-negative breast cancer (18, 19). One possible reason is that radiogenomics can simultaneously provide information regarding macroscopic and microscopic features of the tumor tissue.

Hence, the purpose of this study was to develop and validate a radiogenomics model for the prediction of ALNM in breast cancer compared to a genomics model and a radiomics model.

Materials and methods

Participant characteristics

This study was conducted with patients retrospectively enrolled from a public database and our hospital. The public
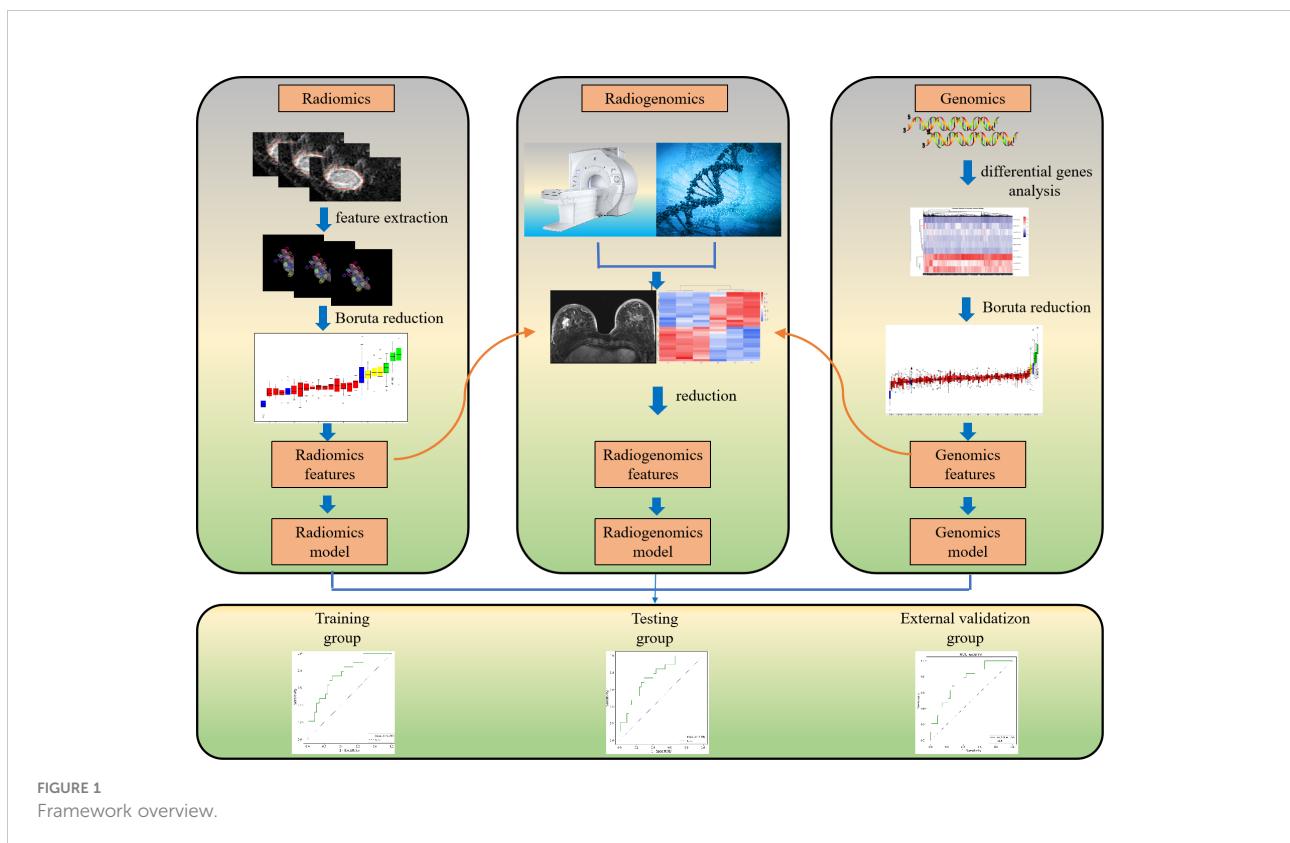
database was originally submitted to The Cancer Genome Atlas (TCGA) and the Cancer Imaging Archive (TCIA) by the contributing institutions under an Institutional Review Board-approved protocol. The second part of the study was approved by the ethics committee of our hospital, and the requirement for individual consent for this retrospective analysis was waived between January 2021 and December 2021. Clinical, breast MRI, and transcriptome data for all patients from TCGA and TCIA, and our hospital were integrated and analyzed. A general overview of the analysis protocol is shown in Figure 1.

The inclusion criteria were as follows: (i) patients had histologically confirmed unilateral primary breast cancer; (ii) patients who underwent breast surgery without neoadjuvant chemotherapy, and sentinel lymph node biopsy or ALN dissection with curative intent; (iii) patients with suspected positive ALN by clinical and/or imaging examinations and plan to receive neoadjuvant therapy, lymph node status was determined by ultrasound-guided core needle biopsy before neoadjuvant therapy (20, 21); (iv) availability of clinical data (age), T staging, complete pathological data for ALN and molecular subtype; (v) breast DCE-MRI was conducted before anti-tumor treatment and core needle biopsy for evaluating tumor extension, intraductal spread, and the presence of multicentric or multifocal lesions; and (vi) availability of transcriptome data. The exclusion criteria were as follows: (i)

insufficient MRI quality to obtain measurements and (ii) patients with multifocal lesions. (iii) patients with negative ALN by clinical and/or imaging examinations and plan to receive neoadjuvant therapy. Details are provided in Figure 2.

Genomics data analysis

Fifteen patients with breast cancer from our hospital were used for the external validation group. Postoperative tumor samples were extracted from the primary tumor sites, and snap-frozen breast cancer samples were collected by the Tumor Samples and Genetic Information Resource Library. Next, the sample from the hospital was processed using RNA extraction and sequencing. Total RNA was extracted from tissue samples with TRIzol reagent (Invitrogen, Carlsbad, CA, USA). RNA purity was monitored on 1% agarose gel. Then, cDNA was synthesized from total RNA using the PrimeScript RT reagent Kit with gDNA Eraser (TaKaRa, RR047A). Then, all patients' transcriptomic data were analyzed for differentially expressed genes (DEGs). The DEGs between the groups with and without ALNM were screened using DESeq (http://bioconductor.org/packages/release/bioc/html/DESeq.html) to detect DEGs with gene symbol annotation with thresholds of $P < 0.05$ and |fold change| > 2.
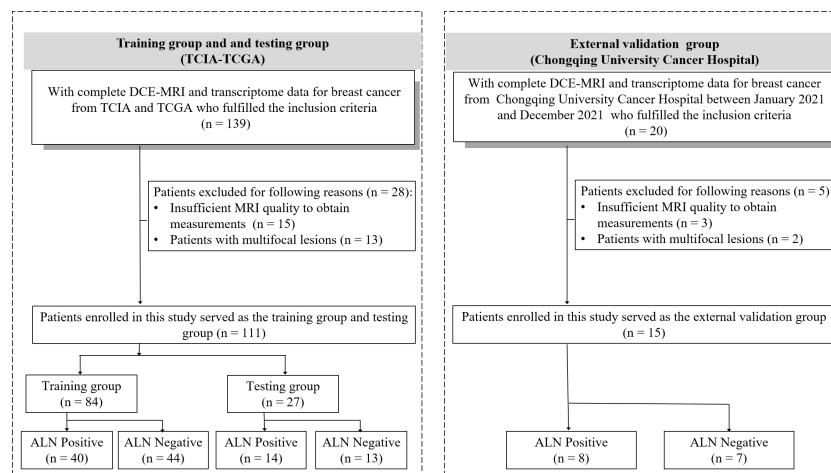

FIGURE 1
Framework overview.

FIGURE 2
Patient recruitment workflow. DCE, dynamic contrast enhanced; TCGA, The Cancer Genome Atlas; TCIA, The Cancer Imaging Archive; ALN, axillary lymph node.

DCE-MRI data acquisition

The public MRI data of breast cancer were downloaded from TCIA. The breast cancer MRI was performed at our hospital using a commercially available 1.5-T system (Philips Ingenia, the Netherlands) with an eight-channel breast array coil. Routine precontrast MRI included T1-weighted imaging, fat-suppressed T2-weighted imaging, and diffusion-weighted imaging. The gadolinium contrast agent (Hengrui, Jiangsu, China) was injected at a rate of 2 mL/s for a dose of 0.1 mmol/kg, followed by a 20 mL saline flush at a rate of 2 mL/s. One precontrast phase image and seven postcontrast phase images were acquired using the axial e-THRIVE polyphase sequence with the following parameters: TR/TE = 5.1/2.5 ms; matrix = 252 × 348; flip angle = 10°; pixel size = 1 × 1 mm; and slice thickness = 2 mm.

Radiomics feature extraction

The regions of interest (ROIs) were delineated manually in each slice of the T1+C data (the peak-enhanced phase of the multiphase contrast-enhanced MRI selected according to the time-intensity curve) by excluding the air, necrosis, and calcification areas *via* the Dr. Wise Multimodal Research Platform (https://keyan.deepwise.com,V1.6.3) (Beijing Deepwise & League of PHD Technology Co., Ltd., Beijing, China). Each DCE-MRI case was reviewed by two experienced radiologists (XXW and LL with 9 and 8 years of experience, respectively). After manual segmentation, consistency intraclass

correlation coefficient (ICC) (22) analysis was performed to measure the features' observer repeatability. ICC scores greater than 0.8 are generally considered to indicate good repeatability. We standardized the image processing procedure according to the image biomarker standardization initiative reference manual (23). Next, 1651 radiomics features were extracted from the normalized image using Pyradiomics (http://www.radiomics.io/pyradiomics.html). Then, they were normalized with Z scores to obtain a standard normal distribution of image intensities. The radiomics features were composed of the following eight types of features: first-order statistics, shape-based (3D), shape-based (2D), gray level cooccurrence matrix (GLCM), gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), neighboring gray-tone difference matrix (NGTDM), gray level dependence matrix (GLDM). All of these features have generally been used in previous radiomics studies (24, 25).

Genomics and radiomics feature selection

We used the Boruta method (R3.6.1 with Boruta version 5.2.0) from a coarse to fine feature reduction strategy to reduce both genomics and radiomics features (24). It was divided into the two steps described below. (i) Univariate analysis was performed using *Student's t* test or the *Mann–Whitney U test* to compare genomics and radiomics features between the groups with and without LNM. All features were ranked in ascending order according to the *P* value, and the top 5% of features were retained for further analysis. (ii) *Spearman* correlation analysis was used to eliminate redundant features. All genomics and

radiomics features with correlation coefficients > 0.85 were detected, and the features with lower *P* values were retained. Then, the selected features were used for further model construction and validation.

Model construction and validation

Patients from the TCIA-TCGA were randomly divided into training and testing group at a ratio of 3:1, and 15 patients from our hospital were enrolled as the external validation group. Logistic regression analysis (Python3.7 with sklearn version 1.1.2) was used to construct radiomics, genomics and radiogenomics models for predicting ALN status. The radiogenomics model was conducted through dimensionality reduction by integrating the features of the radiomics model and the features of the genomics model. All models were trained based on TCGA datasets, and fivefold cross validation was used to determine the parameters of the logistic regression models. We evaluated the prediction performance of the radiomics, genomics and radiogenomics models in the training, testing and external validation groups by constructing receiver operating characteristic (ROC) curves and calculating areas under the curves (AUCs). Decision curve analysis was applied to estimate the clinical utility of the three models. In addition, a calibration curve was generated to evaluate the consistency between the predicted value and the true value.

Statistical analysis

Statistical analyses were performed by commercially available statistical software (SPSS software, version 25.0; Armonk, US). Descriptive statistics were summarized as the means ± standard deviations. Categorical variables were expressed as numbers. Continuous clinical variables were compared using *Student's t* test or the *Mann–Whitney U test.* For categorical variables, *chi-square* tests were used to test differences between groups. The performances of the radiomics model, genomics model and radiogenomics model were compared using the *Delong* test. For all tests, *P* < 0.05 was considered as statistically significant.

TABLE 1 The clinical characteristics of patients in the training, testing and external validation groups.

Characteristics	Training group (n=84)	Testing group (n=27)	External validation group (n=15)	*P* value
Age, mean ± SD, years (range)	53.39 ± 11.44 years (range, 29~80 years)	55.96± 10.56 years (range, 31~82 years)	50.53 ± 7.24 years (range, 42~65years)	0.91
Clinical stage				0.00
I	16	7	3	
II	57	17	7	
III	11	3	5	
T stage				0.00
T1	33	12	4	
T2	45	14	10	
T3	6	1	1	
Axillary lymph node status				0.92
Positive	40	13	8	
Negative	44	14	7	
Molecular subtype				0.00
Luminal A	52	20	3	
Luminal B	11	3	8	
HER2-enriched	6	1	4	
Basel-like	15	3	0	

SD, standard deviation; HER2, human epidermal growth factor receptor 2.

Results

Baseline characteristics

The baseline characteristics in the training, testing, and external validation group are presented in Table 1. A total of 126 patients with breast cancer from one hospital and public database were enrolled in the study. The training group (53.39 ± 11.44 years) included 84 patients, and the testing group (55.96 ± 10.56 years) included 27 patients from the TCGA and TCIA datasets. The external validation group (50.53 ± 7.24 years) included 15 patients who were from our hospital. There were no significant differences in age or LNM among the three groups ($P = 0.91$ and 0.92).

Development and validation of the genomics model

For differential gene expression analysis, 136 DEGs were found between patients with and without ALNM (Figure S1). As shown in Figure S1A, expression levels of 94 genes were increased and 42 genes were decreased between patients with and without ALNM. Heatmap representation of DEGs revealed a similarity between patients with and without ALNM (Figure S1B). Then, 9 genomics features were selected through Boruta analysis (Figure 3A). The 9 genomics features were used to develop the genomics model for predicting ALNM in breast cancer. The AUCs of the training, testing, and external validation group were 0.80, 0.67 and 0.52, respectively (Figure 4A). The AUC and corresponding sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) values of the genomics model are displayed in detail in Table 2. The calibration curves for the

probability of the genomics model for ALNM in the training group, testing group and external validation group are shown in Figures S2A–C.

Development and validation of the radiomics model

A total of 1651 features were extracted from DCE-MRI from every patient. After the reduction by Boruta (Figure 3B), 3 radiomics features were used to construct and validate the radiomics model. The AUCs of the prediction performance of ALNM in breast cancer of the radiomics in the training and testing group were 0.72 and 0.68, respectively (Figure 4B). In addition, the performance of the radiomics model was validated in an independent external validation group, and the AUC of the validation group was 0.71. The AUC and corresponding sensitivity, specificity, PPV and NPV values of the radiomics model are displayed in detail in Table 2. The calibration curves for the probability of the radiomics model for ALNM in the training group, testing group and external validation group are shown in Figures S2D–F.

Development and validation of the radiogenomics model

After stepwise logistic regression with both direction, 5 genomics features and 3 radiomics features were finally selected to develop a radiogenomics model. We performed correlation analysis between the 3 radiomics and 5 genomics features, as shown in Table S1. The results showed that the correlations between the genomics and radiomics features were slightly weaker. The AUCs of predicting ALNM for the
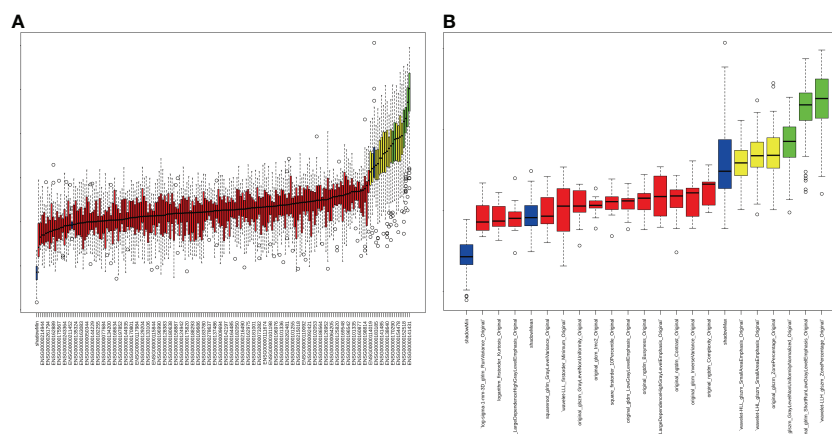

FIGURE 3
Strategy for feature selection using the Boruta method. **(A)** Genomics feature selection; **(B)** Radiomics feature selection.
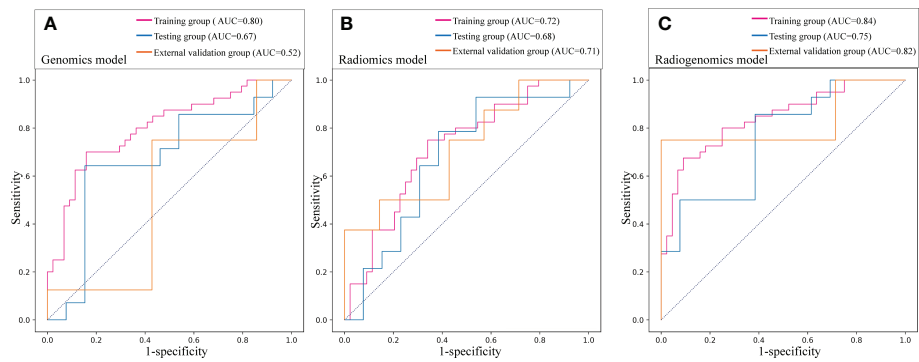
FIGURE 4
Prediction performance of the three models in the training, testing and external validation groups. **(A)** Receiver operating characteristic (ROC) curve of the genomics model; **(B)** ROC curve of the radiomics model; **(C)** ROC curve of the radiogenomics model.

TABLE 2 The performance of three models in the training, testing and external validation groups.

Characteristics	Training group	Testing group	External validation group
Genomics model			
AUC	0.80	0.67	0.52
Sensitivity (%)	70.00	71.43	37.50
Specificity (%)	84.09	46.15	57.14
Accuracy (%)	77.38	59.26	46.67
PPV (%)	80.00	58.83	50.00
NPV (%)	75.51	60.00	44.44
Radiomics model			
AUC	0.72	0.68	0.71
Sensitivity	75.00	78.57	50.00
Specificity	65.91	53.85	57.14
Accuracy	70.24	66.67	53.33
PPV (%)	66.67	64.71	57.14
NPV (%)	74.36	70.00	50.00
Radiogenomics model			
AUC	0.84	0.75	0.82
Sensitivity	67.50	71.43	62.50
Specificity	90.91	61.54	100.00
Accuracy	79.76	66.67	80.00
PPV (%)	87.10	66.67	100.00
NPV (%)	75.47	66.67	70.00

PPV, positive predictive value; NPV, negative predictive value.

radiogenomics model in the training group, testing group and external validation group were 0.84, 0.75 and 0.82, respectively (Figure 4C). The AUC and corresponding sensitivity, specificity, accuracy, PPV, and NPV values of the three groups are detailed in Table 2. Good agreement of the radiogenomics model between the observation and prediction was assessed by the calibration curve, which showed that the bias-corrected line lay close to the ideal curve in the training, testing, and validation group (Figures S2G–I). In addition, the decision curve showed that the radiogenomics model could add more benefit to the prediction of ALNM than the genomics model and radiomics model in the three groups (Figure 5).

The prediction profiles of the radiogenomics model compared with the genomics model and radiomics model by the *DeLong* test in the training group, testing group and external validation group are shown in Figure 6. The results showed that the radiogenomics model significantly improved the AUC compared with the radiomics model in the training group (*P* = 0.01) and showed superior performance compared with the genomics model in the external validation group (*P* = 0.02) (Table 3).

Discussion

In this study, we established three models for the assessment of ALNM in patients with breast cancer. The radiogenomics model incorporating quantitative radiomics and genomics features could accurately predict LNM with favorable AUC, high specificity, and PPV, which were superior to those of the radiomics and genomics model. The robustness and generalizability of the radiogenomics model were further tested in a multicohort and validated in one hospital.

Preoperative LNM prediction could be beneficial for breast cancer patients. Previous studies have identified that genomics and epigenomic markers unraveled significant epigenetic

changes during the progression from primary breast tumor to LNM, which may contribute to improved prognosis and prediction in breast cancer (26). The gene expression data could predict LNM, and the value of such patterns resulted in a predictive accuracy of approximately 90% (27). In this study, we developed a genomics model consisting of 9 features with good performance in predicting LNM. These genes included PTPN21 (28), ST6GALNAC3 (29), FAM13A (30), and CHRNA7 (31), which were correlated with metastasis. These genes were also enriched in the PI3K/Akt, Notch1/Hes1 or Akt1/mTOR signaling pathways and were reported to be correlated with LNM in breast cancer (32–35). The NMRK2 gene was identified as an important target for mitochondrial respiration, and mitochondrial respiration is frequently dependent on metastatic cells (36, 37). A study showed that the recurrence-free survival of patients with loss of function of ZFP36L2 was significantly shorter than that of patients with no loss of ZFP36L2 function in colorectal cancer (38).

The results of our study confirmed that the radiomics model could noninvasively predict ALNM in breast cancer. Previous research has shown that the radiomics model alone predicted LNM with AUCs of 0.76 (39) and 0.806 (40) in breast cancer patients, and their prediction ability was moderate. When the above radiomics features were combined with clinical features or clinicopathologic characteristics, the AUC improved significantly. However, our present study enrolled patients from TCIA, which is a multiagency mixed data, and the radiomics features were not combined with clinical features; therefore, the AUC values of the radiomics model were lower than those of previous studies.

Radiogenomics is an emerging field of cancer research. The first study on radiogenomics in breast cancer was published in 2012 (16), which revealed that radiogenomics analysis of breast cancer with MRI is a novel method that can be used to understand the underlying molecular biology of breast cancers. Then, an increasing number of studies have mainly focused on exploring the correlation between the morphological and
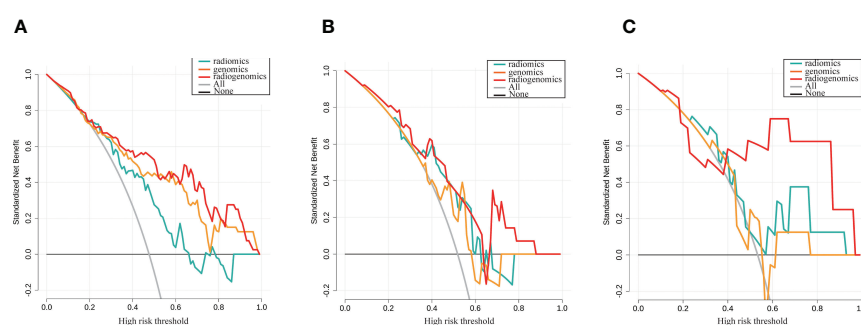

FIGURE 5
Decision curve analysis for the genomics, radiomics and radiogenomics models of LNM prediction in breast cancer. **(A)** Training group; **(B)** Testing group; **(C)** External validation group. The y-axis measures the standardized net benefit.
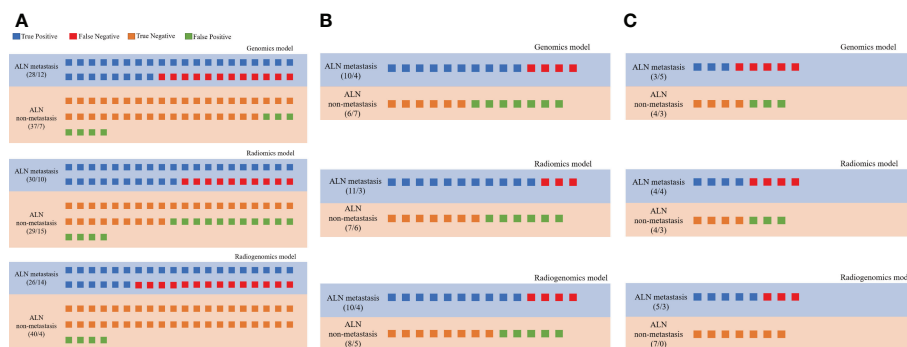
FIGURE 6
Prediction profiles of the three models. All individual participants were identified as "LNM" or "non-LNM", and recognized as true positive (blue), false negative (red), true negative (orange) and false positive (green) according to their true labels of LNM in all groups. **(A)** Prediction profiles of the training group; **(B)** Prediction profiles of the testing group; **(C)** Prediction profiles of the external validation groups.

enhancement features of DCE-MRI and the genomics features of molecular subtypes (41, 42). In our results, the radiogenomics model had a higher predictive performance than the radiomics-only model or the genomics-only model to predict ALNM in breast cancer because it incorporated macroscopic genomics features and high-throughput radiological features to enhance the predictive value and discover novel biomarkers. Our results showed that genomics features have few associations with MRI-derived radiomics features, which suggests that genes provide additional information for structural imaging. Radiogenomics can effectively predict LNM by bridging the limitations of genomics and radiomics and assisting clinicians in making more precise clinical decisions.

The 2022 NCCN Clinical Practice Guidelines in Oncology recommend that ultrasound-guided fine-needle aspiration cytology or core needle biopsy can be performed for patients with lymph nodes suspected to be positive by clinical and/or imaging examinations or patients considering systemic therapy before surgery (43). Previous study (44) confirmed that ultrasound-guided core needle biopsy was superior to ultrasound-guided fine-needle aspiration cytology in diagnosing axillary nodal metastases: sensitivity 88% (95 confidence interval: 84% to 91%) versus 74% (95 confidence

interval: 70% to 78%) respectively, and they both a high specificity of 100%. In our study, we used US guided core needle biopsy in diagnosing ALN metastasis in patients who received neoadjuvant chemotherapy. Therefore, this ensures the reliability and accuracy of this study.

Our study had some limitations. First, this was a retrospective study, and a larger sample size would be desirable. Given the lack of multicenter data, we compiled the training and testing groups from the TCIA-TCGA and a verification group from our hospital database, thereby extending the generalizability of the model. Second, we only focused our prediction on DCE without considering other MRI techniques, such as T2-weighted imaging and DWI. Finally, only logistic regression analysis was used to construct the prediction model, and future studies should consider deep learning or more machine learning algorithms.

By integrating radiomics and genomics features, we built a radiogenomics prediction model that can significantly improve the performance to predict ALNM in breast cancer. The radiogenomics prediction model might reduce unnecessary ALN dissection and improve the quality of life of cancer patients, which could contribute to the realization of precision medicine in breast cancer.

TABLE 3 Delong test for prediction improvements of radiogenomics model in the training, testing and external validation groups.

Characteristics	Training group	P value	Testing group	P value	External validation group	P value
Radiomics model vs. Genomics model						
DeLong test	-0.08	0.27	0.01	0.97	0.19	0.41
radiogenomics model vs. Radiomics model						
DeLong test	0.12	**0.01**	0.07	0.43	0.11	0.50
radiogenomics model vs. Genomics model						
DeLong test	0.04	0.28	0.08	0.35	0.30	**0.02**
The bold values represent a p-value less than 0.05, indicating a significant improvement in prediction.						

Data availability statement

Ethics statement

The studies involving human participants were reviewed and approved by The public database was originally submitted to The Cancer Genome Atlas (TCGA) and the Cancer Imaging Archive (TCIA) by the contributing institutions under an Institutional Review Board-approved protocol. The second part of the study was approved by the ethics committee of our hospital(CZLS2022030-A). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

Conceptualization, HC and XL. Methodology, XL, FJ and LW. Validation, TY and LL. Data Analysis, HC, XL, SL and ST. Writing – Original Draft Preparation, HC. Writing – Review & Editing, XW and JZ. Supervision, YH, YC, WW. Funding Acquisition, HC, XW and JZ. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.1076267/full#supplementary-material

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* (2020) 70(1):7–30. doi: 10.3322/caac.21590

2. Maxwell F, de Margerie Mellon C, Bricout M, Cauderlier E, Chapelier M, Albiter M, et al. Diagnostic strategy for the assessment of axillary lymph node status in breast cancer. *Diagn Interv Imaging* (2015) 96(10):1089–101. doi: 10.1016/j.diii.2015.07.007

3. Lyman GH, Somerfield MR, Bosserman LD, Perkins CL, Weaver DL, Giuliano AE. Sentinel lymph node biopsy for patients with early-stage breast cancer: American society of clinical oncology clinical practice guideline update. *J Clin Oncol* (2017) 35(5):561–4. doi: 10.1200/JCO.2016.71.0947

4. Boughey JC, Moriarty JP, Degnim AC, Gregg MS, Egginton JS, Long KH. Cost modeling of preoperative axillary ultrasound and fine-needle aspiration to guide surgery for invasive breast cancer. *Ann Surg Oncol* (2010) 17(4):953–8. doi: 10.1245/s10434-010-0919-1

5. Shaitelman SF, Cromwell KD, Rasmussen JC, Stout NL, Armer JM, Lasinski BB, et al. Recent progress in the treatment and prevention of cancer-related lymphedema. *CA Cancer J Clin* (2015) 65(1):55–81. doi: 10.3322/caac.21253

6. Expert Panel on Breast, I, Slanetz PJ, Moy L, Baron P, diFlorio RM, Green ED, et al. ACR appropriateness Criteria((R)) monitoring response to neoadjuvant systemic therapy for breast cancer. *J Am Coll Radiol* (2017) 14(11S):S462–75. doi: 10.1016/j.jacr.2017.08.037

7. Zheng X, Yao Z, Huang Y, Yu Y, Wang Y, Liu Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun* (2020) 11(1):1236. doi: 10.1038/s41467-020-15027-z

8. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14(12):749–62. doi: 10.1038/nrclinonc.2017.141

9. Song D, Yang F, Zhang Y, Guo Y, Qu Y, Zhang X, et al. Dynamic contrast-enhanced MRI radiomics nomogram for predicting axillary lymph node metastasis in breast cancer. *Cancer Imaging* (2022) 22(1):17. doi: 10.1186/s40644-022-00450-w

10. Qiu X, Fu Y, Ye Y, Wang Z, Cao C. A nomogram based on molecular biomarkers and radiomics to predict lymph node metastasis in breast cancer. *Front Oncol* (2022) 12:790076. doi: 10.3389/fonc.2022.790076

11. Ma M, Jiang Y, Qin N, Zhang X, Zhang Y, Wang X, et al. A radiomics model for preoperative predicting sentinel lymph node metastasis in breast cancer based on dynamic contrast-enhanced MRI. *Front Oncol* (2022) 12:884599. doi: 10.3389/fonc.2022.884599

12. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* (2005) 365(9460):671–9. doi: 10.1016/S0140-6736(05)17947-1

13. Shang C, Li Y, He T, Liao Y, Du Q, Wang P, et al. The prognostic miR-532-5p-correlated ceRNA-mediated lipid droplet accumulation drives nodal metastasis of cervical cancer. *J Adv Res* (2022) 37:169–84. doi: 10.1016/j.jare.2021.09.009

14. Orozco JIJ, Le J, Ensenyat-Mendez M, Baker JL, Weidhaas J, Klomhaus A, et al. Machine learning-based epigenetic classifiers for axillary staging of patients with ER-positive early-stage breast cancer. *Ann Surg Oncol* (2022) 29(10):6407–14. doi: 10.1245/s10434-022-12143-6

15. Grimm LJ, Mazurowski MA. Breast cancer radiogenomics: Current status and future directions. *Acad Radiol* (2020) 27(1):39–46. doi: 10.1016/j.acra.2019.09.012

16. Yamamoto S, Maki DD, Korn RL, Kuo MD. Radiogenomic analysis of breast cancer using MRI: a preliminary study to define the landscape. *AJR Am J Roentgenol* (2012) 199(3):654–63. doi: 10.2214/AJR.11.7824

17. Yamamoto S, Han W, Kim Y, Du L, Jamshidi N, Huang D, et al. Breast cancer: Radiogenomic biomarker reveals associations among dynamic contrast-enhanced MR imaging, long noncoding RNA, and metastasis. *Radiology* (2015) 275(2):384–92. doi: 10.1148/radiol.15142698

18. Guo W, Li H, Zhu Y, Lan L, Yang S, Drukker K, et al. Prediction of clinical phenotypes in invasive breast carcinomas from the integration of radiomics and genomics data. *J Med Imaging (Bellingham)* (2015) 2(4):41007. doi: 10.1117/1.JMI.2.4.041007

19. Zhang Y, You C, Pei Y, Yang F, Li D, Jiang YZ, et al. Integration of radiogenomic features for early prediction of pathological complete response in patients with triple-negative breast cancer and identification of potential therapeutic targets. *J Transl Med* (2022) 20(1):256. doi: 10.1186/s12967-022-03452-1

20. Gennari A, Andre F, Barrios CH, Cortes J, de Azambuja E, DeMichele A, et al. ESMO clinical practice guideline for the diagnosis, staging and treatment of patients with metastatic breast cancer. *Ann Oncol* (2021) 32(12):1475–95. doi: 10.1016/j.annonc.2021.09.019

21. Brackstone M, Baldassarre FG, Perera FE, Cil T, Chavez Mac Gregor M, Dayes IS, et al. Management of the axilla in early-stage breast cancer: Ontario health (Cancer care Ontario) and ASCO guideline. *J Clin Oncol* (2021) 39(27):3056–82. doi: 10.1200/JCO.21.00934

22. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* (2016) 15(2):155–63. doi: 10.1016/j.jcm.2016.02.012

23. Zwanenburg A, Vallières M, Abdalah M, Aerts H, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145

24. Liu Z, Li Z, Qu J, Zhang R, Zhou X, Li L, et al. Radiomics of multiparametric MRI for pretreatment prediction of pathologic complete response to neoadjuvant chemotherapy in breast cancer: A multicenter study. *Clin Cancer Res* (2019) 25(12):3538–47. doi: 10.1158/1078-0432.Ccr-18-3190

25. Huang Y, Liang C, He L, Tian J, Liang C, Chen X, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J Clin Oncol* (2016) 34(18):2157–64. doi: 10.1200/jco.2015.65.9128

26. Urrutia G, Laurito S, Marzese DM, Gago F, Orozco J, Tello O, et al. Epigenetic variations in breast cancer progression to lymph node metastasis. *Clin Exp Metastasis* (2015) 32(2):99–110. doi: 10.1007/s10585-015-9695-4

27. Huang E, Cheng SH, Dressman H, Pittman J, Tsou MH, Horng CF, et al. Gene expression predictors of breast cancer outcomes. *Lancet* (2003) 361(9369):1590–6. doi: 10.1016/S0140-6736(03)13308-9

28. Li XQ, Liu BC, Jiang XB, Gao WW, Rong B, Wei Y, et al. Inhibition of PTPN21 has antitumor effects in glioma by restraining the EGFR/PI3K/AKT pathway. *Toxicol Appl Pharmacol* (2022) 451:116180. doi: 10.1016/j.taap.2022.116180

29. Yuan Q, Chen X, Han Y, Lei T, Wu Q, Yu X, et al. Modification of alpha2,6-sialylation mediates the invasiveness and tumorigenicity of non-small cell lung cancer cells in vitro and *in vivo via* Notch1/Hes1/MMPs pathway. *Int J Cancer* (2018) 143(9):2319–30. doi: 10.1002/ijc.31737

30. Li R, Tan S, Yu M, Jundt MC, Zhang S, Wu M. Annexin A2 regulates autophagy in pseudomonas aeruginosa infection through the Akt1-mTOR-ULK1/2 signaling pathway. *J Immunol* (2015) 195(8):3901–11. doi: 10.4049/jimmunol.1500967

31. Xiang T, Yu F, Fei R, Qian J, Chen W. CHRNA7 inhibits cell invasion and metastasis of LoVo human colorectal cancer cells through PI3K/Akt signaling. *Oncol Rep* (2016) 35(2):999–1005. doi: 10.3892/or.2015.4462

32. Wang LL, Hao S, Zhang S, Guo LJ, Hu CY, Zhang G, et al. PTEN/PI3K/AKT protein expression is related to clinicopathological features and prognosis in breast cancer with axillary lymph node metastases. *Hum Pathol* (2017) 61:49–57. doi: 10.1016/j.humpath.2016.07.040

33. Mohammed SI, Torres-Luquis O, Zhou W, Lanman NA, Espina V, Liotta L. Tumor-draining lymph secretome en route to the regional lymph node in breast cancer metastasis. *Breast Cancer (Dove Med Press)* (2020) 12:57–67. doi: 10.2147/BCTT.S236168

34. Yang B, Huang CZ, Yu T, Zhou SN, Liu Q, Liu GJ, et al. Metformin depresses overactivated Notch1/Hes1 signaling in colorectal cancer patients with type 2 diabetes mellitus. *Anticancer Drugs* (2017) 28(5):531–9. doi: 10.1097/CAD.0000000000000483

35. Ren C, Pan R, Hou L, Wu H, Sun J, Zhang W, et al. Suppression of CLEC3A inhibits osteosarcoma cell proliferation and promotes their chemosensitivity through the AKT1/mTOR/HIF1alpha signaling pathway. *Mol Med Rep* (2020) 21(4):1739–48. doi: 10.3892/mmr.2020.10986

36. Chen Y, Yang L, Lu Y, Liu N, Ma W, Fan H, et al. Up-regulation of NMRK2 mediated by TFE3 fusions is the key for energy metabolism adaption of Xp11.2 translocation renal cell carcinoma. *Cancer Lett* (2022) 538:215689. doi: 10.1016/j.canlet.2022.215689

37. Passaniti A, Kim MS, Polster BM, Shapiro P. Targeting mitochondrial metabolism for metastatic cancer therapy. *Mol Carcinog* (2022) 61(9):827–38. doi: 10.1002/mc.23436

38. Kishigami F, Tanaka Y, Yamamoto Y, Ueno T, Kojima S, Sato K, et al. Exploration of predictive biomarkers for postoperative recurrence of stage II/III colorectal cancer using genomic sequencing. *Cancer Med* (2022) 11(18):3457-3470. doi: 10.1002/cam4.4710

39. Han L, Zhu Y, Liu Z, Yu T, He C, Jiang W, et al. Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer. *Eur Radiol* (2019) 29(7):3820–9. doi: 10.1007/s00330-018-5981-2

40. Liu C, Ding J, Spuhler K, Gao Y, Serrano Sosa M, Moriarty M, et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer by radiomic signatures from dynamic contrast-enhanced MRI. *J Magn Reson Imaging* (2019) 49(1):131–40. doi: 10.1002/jmri.26224

41. Mazurowski MA, Zhang J, Grimm LJ, Yoon SC, Silber JI. Radiogenomic analysis of breast cancer: luminal b molecular subtype is associated with enhancement dynamics at MR imaging. *Radiology* (2014) 273(2):365–72. doi: 10.1148/radiol.14132641

42. Sutton EJ, Dashevsky BZ, Oh JH, Veeraraghavan H, Apte AP, Thakur SB, et al. Breast cancer molecular subtype classifier that incorporates MRI features. *J Magn Reson Imaging* (2016) 44(1):122–9. doi: 10.1002/jmri.25119

43. Gradishar WJ, Moran MS, Abraham J, Aft R, Agnese D, Allison KH, et al. Breast cancer, version 3.2022, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw* (2022) 20(6):691–722. doi: 10.6004/jnccn.2022.0030

44. Balasubramanian I, Fleming CA, Corrigan MA, Redmond HP, Kerin MJ, Lowery AJ. Meta-analysis of the diagnostic accuracy of ultrasound-guided fine-needle aspiration and core needle biopsy in diagnosing axillary lymph node metastasis. *Br J Surg* (2018) 105(10):1244–53. doi: 10.1002/bjs.10920